1600
-1652 #5

```
RAW SEQUENCE LISTING                    DATE: 11/28/2001
PATENT APPLICATION:  US/09/840,762A     TIME: 12:17:40

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw
```

ENTERED

```
 1 <110> APPLICANT: Vreeland, Valerie
 2       Ng, Kwan L.
 3       The Regents of the University of California
 4 <120> TITLE OF INVENTION: Recombinant Vanadium Haloperoxidases and Their Uses
 5 <130> FILE REFERENCE: 023070-087100US
 6 <140> CURRENT APPLICATION NUMBER: 09/840,762A
 7 <141> CURRENT FILING DATE: 2001-04-23
 8 <150> PRIOR APPLICATION NUMBER: 09/151,189
 9 <151> PRIOR FILING DATE: 2001-04-23
10 <160> NUMBER OF SEQ ID NOS: 11
11 <170> SOFTWARE: PatentIn Ver. 2.0
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 2931
15 <212> TYPE: DNA
16 <213> ORGANISM: Fucus distichus
17 <220> FEATURE:
18 <221> NAME/KEY: CDS
19 <222> LOCATION: (228)..(2258)
20 <223> OTHER INFORMATION: vanadium bromoperoxidase
21 <400> SEQUENCE: 1
22       cgcggacaag ccttggaaga gaggttgccc aattcaacag agcgaggccc gtgaaggtgt 60
23       ggaggacacg tgctacaagc tgatccacga gagcctcaac ttccctactg atacgggagt 120
24       ttgtactgcg ccgcgttgcc aaaaaccgca actttaaaca gcgctcgcga gcgccacatg 180
25       cttcccacgc atccacaaaa tcgacagtgg tatcgctgag cttgaat atg ctt tgc 236
26                                                       Met Leu Cys
27                                                        1
28       cat gca gcg gac acg aca aga ggc tct cct atg cct gac acc gga gtg 284
29       His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp Thr Gly Val
30            5                   10                  15
31       ctt cgg ttg ctc aca tca gag cag cgc gct aaa ggt tgg aga cgc cag 332
32       Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp Arg Arg Gln
33        20                  25                  30                  35
34       tta gag ggg gag aaa tca cta ggt ttt cat cca agc gag acg cct tat 380
35       Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu Thr Pro Tyr
36                        40                  45                  50
37       atc aag tac ttg gaa ggc tct gag act tgg aag aag gtt aag ctt cca 428
38       Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val Lys Leu Pro
39                    55                  60                  65
40       acg gac ggc ata tcg gct tcc aag atc ctg ggt aaa att atg gcc agg 476
41       Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile Met Ala Arg
42                    70                  75                  80
43       gtc cgc atc gct acc gcc ttg gct gtg gta ctg gcc gca ccc tgt ttg 524
44       Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala Pro Cys Leu
45            85                  90                  95
46       gca ttc gac gag gtc aca gcc agt ggt gtt ttc cct gag gaa cac aag 572
47       Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu Glu His Lys
48       100                 105                 110                 115
```

RAW SEQUENCE LISTING   DATE: 11/28/2001
PATENT APPLICATION: US/09/840,762A   TIME: 12:17:40

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw

```
49   cac acc ggg gag gga aga cac ctc cag acc tgt aca aac tcc gac gat
50   His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn Ser Asp Asp
51             120                 125                 130              668
52   gcg ctg gat ccg acg gcg ccg aat aga agg gac aac gta gct ttt gcg
53   Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val Ala Phe Ala
54             135                 140                 145
55   tcg cgg cgc gat gcc gcc agg cga gaa cgt gac ggg aca ggg act gtc   716
56   Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr Gly Thr Val
57             150                 155                 160
58   tgc caa atc act aac gga gaa act gat ttg gct acc atg ttc cac aag   764
59   Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met Phe His Lys
60             165                 170                 175
61   tct ctg cca cac gat gaa ctg gga cag gta acc gca gac gac ttc gct   812
62   Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp Asp Phe Ala
63   180                 185                 190                 195
64   atc ctc gag gac tgc atc tta aac gga gat ttc agc att tgc gag gac   860
65   Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile Cys Glu Asp
66             200                 205                 210
67   gtg cct gcg gga gac ccg gcg ggt cgc ctc gtc aat cct acc gct gcg   908
68   Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro Thr Ala Ala
69             215                 220                 225
70   ttt gcc atc gac ata tcc ggt ccc gca ttc tcg gct acg aca ata ccc   956
71   Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr Thr Ile Pro
72             230                 235                 240
73   ccg gta cct acc ctt tcc tct cct gag ctc gcc gct cag ttg gcg gag   1004
74   Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln Leu Ala Glu
75             245                 250                 255
76   cta tac tgg atg gcg ctg gcc agg gat gta ccc ttt atg cag tat ggc   1052
77   Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met Gln Tyr Gly
78   260                 265                 270                 275
79   acc gac gaa att acc act acc gcg gca gcc aac ctc gct gga atg gga   1100
80   Thr Asp Glu Ile Thr Thr Thr Ala Ala Ala Asn Leu Ala Gly Met Gly
81             280                 285                 290
82   ggc ttc cca aat ctg gac gcc gtg tcg ata ggg tcc gat ggt acg gtg   1148
83   Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp Gly Thr Val
84             295                 300                 305
85   gac ccg ttc tcc cag ctc ttc cga gcg acc ttc gtt ggt gtt gaa acg   1196
86   Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly Val Glu Thr
87             310                 315                 320
88   ggg ccc ttt gtc tct cag ctg ctc gtg aac agc ttc acc atc gac gct   1244
89   Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr Ile Asp Ala
90   325                 330                 335
91   att acg gtc gaa ccg aag cag gag aca ttc gcc ccc gac ttg aac tat   1292
92   Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp Leu Asn Tyr
93   340                 345                 350                 355
94   atg gtc gat ttt gac gaa tgg ctg aac att cag aat ggt gga ccc ccg   1340
95   Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly Gly Pro Pro
96             360                 365                 370
97   gcc ggc ccc gaa gag tta gac gaa gag ctg cgt ttt atc cgt aac gcc   1388
```

RAW SEQUENCE LISTING   DATE: 11/28/2001
PATENT APPLICATION: US/09/840,762A   TIME: 12:17:40

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw

```
98      Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile Arg Asn Ala
99                375                  380                  385
100     cgc gac ctg gcc agg gtc tcc ttc gtg gac aat atc aac acc gaa gct   1436
101     Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn Thr Glu Ala
102              390                  395                  400
103     tat cgc ggg tct ctt atc cta ctt gag ctg gga gcc ttc agc agg ccc   1484
104     Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe Ser Arg Pro
105          405                  410                  415
106     ggt atc aac ggt cca ttc atc gac agt gat cgg cag gcg ggc ttc gtc   1532
107     Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala Gly Phe Val
108     420                  425                  430                  435
109     aac ttc ggc acg tct cac tac ttc aga ttg ata ggt gcc gcc gag ctg   1580
110     Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala Ala Glu Leu
111              440                  445                  450
112     gcg cag cgt gcc tcg tgt tac caa aag tgg cag gtg cat cga ttt gca   1628
113     Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His Arg Phe Ala
114              455                  460                  465
115     cgc ccc gag gct ctc ggg ggt acc ctc cac aac acc atc gcg ggg gat   1676
116     Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile Ala Gly Asp
117              470                  475                  480
118     cta gat gca gac ttc gac atc tcc ctt ctt gaa aat gat gag ctc ttg   1724
119     Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp Glu Leu Leu
120          485                  490                  495
121     aaa cgt gtg gcg gag ata aat gcg gcg cag aat ccc aac aac gag gtc   1772
122     Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn Asn Glu Val
123     500                  505                  510                  515
124     acc tac ctt ctt cca caa gct atc caa gtg gga tcg cca acg cac cct   1820
125     Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro Thr His Pro
126              520                  525                  530
127     tcc tac ccg tcc ggc cac gct acc caa aat gga gca ttt gcc aca gtt   1868
128     Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe Ala Thr Val
129              535                  540                  545
130     ctg aag gcc ctc att ggc cta gat cgg gga ggt gag tgc ttc cct aac   1916
131     Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys Phe Pro Asn
132              550                  555                  560
133     ccc gtg ttc cca agc gat gac ggc ctg gaa cta atc aac ttc gaa ggg   1964
134     Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn Phe Glu Gly
135          565                  570                  575
136     gca tgc ctt aca tat gag gga gag atc aac aag ctc gcg gtc aac gtc   2012
137     Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala Val Asn Val
138     580                  585                  590                  595
139     gca ttt ggg agg cag atg ctg ggc atc cac tat cgg ttc gac ggt atc   2060
140     Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe Asp Gly Ile
141              600                  605                  610
142     caa ggc cta ctt ctc gga gag aca atc act gta cga aca ctt cac cag   2108
143     Gln Gly Leu Leu Leu Gly Glu Thr Ile Thr Val Arg Thr Leu His Gln
144              615                  620                  625
145     gag ctg atg acg ttc gcc gag gaa gcc acc ttt gaa ttc cgc tta ttc   2156
146     Glu Leu Met Thr Phe Ala Glu Glu Ala Thr Phe Glu Phe Arg Leu Phe
```

RAW SEQUENCE LISTING   DATE: 11/28/2001
PATENT APPLICATION: US/09/840,762A   TIME: 12:17:40

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw

```
147              630           635               640
148    acc gga gag gtc atc aaa ctt ttc cag gac ggg aca ttc tcc atc gat   2204
149    Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe Ser Ile Asp
150      645             650              655
151    gga gat atg tgt tcc ggt ttg gtt tac act ggc gtg gcg gac tgc cag   2252
152    Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala Asp Cys Gln
153    660              665              670               675
154    gct tagtgcagaa aataataatt gtcggatgct taaaatgcac ccacgaccaa        2305
155    Ala
157    gtcgtcgagt cacgtcgccg gagcatcctt cagcgaaaaa ggagagtaac ctatatgcta 2365
158    tagaggagaa ccacggagta caatgcaggt tcttttacca tgtacattgg attgcagtaa 2425
159    gtgcggttag agagggatac gttaaacgtg cttgcctgtg tatatgatac atttgtcatg 2485
160    gaaatattag aatgcgttga cttgacttca ccatgaaata ccatgatcgc gtggtgtgct 2545
161    gctttcacct gtcggagcgg tacgtaagat gtgctttcta ctgagccgtt tgtgtttagt 2605
162    ccattccgcg tggcagtgta aacaaagagg atgtagtctc gccctcagtt tggagagtac 2665
163    cgtaggtggc aggacgtata tctctggtag cggtctgtta agaacttcca caagaccgtt 2725
164    tacgtttggt tgtttagtcg atgcctcttc gttacttgac cgatccattg agagtacctg 2785
165    taccagtatg gtgtaagaca tatttttctc ctgttatgga tctgtagaac agctaggtgt 2845
166    tgttttatac acaggatgct ataaaatagg gatgttgata atggcatcgg tactcatgaa 2905
167    accgcaaaat ggcgatagat attccc                                      2931
```

169 <210> SEQ ID NO: 2
170 <211> LENGTH: 676
171 <212> TYPE: PRT
172 <213> ORGANISM: Fucus distichus
173 <400> SEQUENCE: 2

```
174    Met Leu Cys His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp
175      1              5                10                15
176    Thr Gly Val Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp
177              20              25                 30
178    Arg Arg Gln Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu
179              35              40                 45
180    Thr Pro Tyr Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val
181      50              55                 60
182    Lys Leu Pro Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile
183      65              70                 75                80
184    Met Ala Arg Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala
185              85                 90                 95
186    Pro Cys Leu Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu
187              100             105                110
188    Glu His Lys His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn
189          115             120                125
190    Ser Asp Asp Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val
191        130             135                140
192    Ala Phe Ala Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr
193      145             150                155                160
194    Gly Thr Val Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met
195              165                170                175
196    Phe His Lys Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp
197              180                185                190
```

RAW SEQUENCE LISTING   DATE: 11/28/2001
PATENT APPLICATION: US/09/840,762A   TIME: 12:17:40

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw

```
198   Asp Phe Ala Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile
199              195                 200                 205
200   Cys Glu Asp Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro
201         210                 215                 220
202   Thr Ala Ala Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr
203   225                 230                 235                 240
204   Thr Ile Pro Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln
205                 245                 250                 255
206   Leu Ala Glu Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met
207             260                 265                 270
208   Gln Tyr Gly Thr Asp Glu Ile Thr Thr Thr Ala Ala Ala Asn Leu Ala
209             275                 280                 285
210   Gly Met Gly Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp
211         290                 295                 300
212   Gly Thr Val Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly
213   305                 310                 315                 320
214   Val Glu Thr Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr
215                 325                 330                 335
216   Ile Asp Ala Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp
217             340                 345                 350
218   Leu Asn Tyr Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly
219             355                 360                 365
220   Gly Pro Pro Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile
221         370                 375                 380
222   Arg Asn Ala Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn
223   385                 390                 395                 400
224   Thr Glu Ala Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe
225                 405                 410                 415
226   Ser Arg Pro Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala
227             420                 425                 430
228   Gly Phe Val Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala
229             435                 440                 445
230   Ala Glu Leu Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His
231         450                 455                 460
232   Arg Phe Ala Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile
233   465                 470                 475                 480
234   Ala Gly Asp Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp
235                 485                 490                 495
236   Glu Leu Leu Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn
237             500                 505                 510
238   Asn Glu Val Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro
239             515                 520                 525
240   Thr His Pro Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe
241         530                 535                 540
242   Ala Thr Val Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys
243   545                 550                 555                 560
244   Phe Pro Asn Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn
245                 565                 570                 575
246   Phe Glu Gly Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala
```

VERIFICATION SUMMARY   DATE: 11/28/2001
PATENT APPLICATION: US/09/840,762A   TIME: 12:17:41

Input Set : N:\Crf3\RULE60\09840762A.RAW
Output Set: N:\CRF3\11282001\I840762A.raw